

УДК 621.313.17

И.П. Плисс, к. т. н., ст. науч. сотр.,
С.В. Попов, к. т. н., ст. науч. сотр.,
Т.В. Рыбальченко,
Е.В. Слипченко, к. т. н.
 Харьковский национальный
 университет радиоэлектроники

МОДИФИЦИРОВАННАЯ РЕКУРРЕНТНАЯ НЕЙРОННАЯ СЕТЬ ЭЛМАНА В ЗАДАЧЕ КРАТКОСРОЧНОГО ПРОГНОЗИРОВАНИЯ ПОТРЕБЛЕНИЯ ЭЛЕКТРОЭНЕРГИИ

Введение. В условиях глобального экономического кризиса, усугубленного собственными энергетическими проблемами страны, задача достаточно точного прогнозирования потребления электроэнергии приобретает очень важное значение для решения проблем, связанных с планированием работы энергосистем, распределением генерируемых мощностей, оперативным управлением энергетической отраслью народного хозяйства в целом. Неточное прогнозирование потребления электроэнергии ведет к существенным экономическим потерям. Так, в [1, 2] отмечалось, что для экономики Великобритании за период 1985-2000г. повышение точности прогнозирования на 1% могло бы привести к дополнительным прибылям энергокомпаний порядка 100 миллионов фунтов стерлингов в год. Это связано с тем, что переоценка будущего потребления ведет к неоправданному перерасходу топлива всех видов, а его недооценка – к снижению качества энергоснабжения потребителей. В связи с этим повышение точности прогнозирования потребления электроэнергии действительно является актуальной проблемой, требующей эффективных методов ее решения.

Среди множества подходов, с той или с иной степенью успеха использованных в этой задаче, можно выделить [3] традиционные методы анализа временных рядов и, прежде всего, подход Бокса-Дженкинса, аппарат калмановской фильтрации и теории адаптивных систем, регрессионные, корреляционные и спектральные алгоритмы, и как наиболее эффективные – методы вычислительного интеллекта и, прежде всего, искусственные нейронные сети (ИНС). Успех ИНС в рассматриваемой проблеме объясняется нелинейным характером прогнозируемых процессов, высоким уровнем неопределенности (структурной и параметрической) о внутреннем их характере, стохастичностью и хаотичностью, не связанными с традиционной гипотезой о гауссовости анализируемых сигналов, в общем, со всем тем, что не позволяет эффективно использовать традиционные методы статистического анализа и адаптивного прогнозирования.

На сегодня можно отметить достаточно большое число удачных примеров использования ИНС в задаче прогнозирования потребления электроэнергии как у нас в стране [4-6], так и за рубежом [7-11], при этом интересно отметить, что в подавляющем большинстве случаев в основе прогнозирующей системы лежит, по удачному определению авторов [3], – «рабочая лошадка нейронных сетей» – многослойный персептрон со всеми своими модификациями, объединенными общей архитектурой с прямой передачей информации. Все эти ИНС с точки зрения теории случайных процессов объединяются общим понятием «нелинейная модель авторегрессии с экзогенными входами» (NARX-модель), которая является частным случаем более общих структур, содержащих компоненты скользящего среднего и известных как NARMAX-модели, обладающие большей гибкостью и потенци-

ально более высокой точностью. Теоретически прогнозирующая NARMAX-модель может быть построена на основе обычного многослойного персептрона, охваченного глобальной обратной связью [12], однако обучение такой системы характеризуется низкой скоростью сходимости, необходимостью использования настройки по эпохам, невозможностью работы с нестационарными сигналами. Альтернативой нейронным сетям с прямой передачей информации в задачах прогнозирования могут служить рекуррентные нейронные сети [13], включающие в свою архитектуру как глобальные, так и локальные (на уровне слоев) обратные связи и обучаемые с помощью специализированных процедур. На сегодня известен ряд примеров использования рекуррентных ИНС для решения задачи прогнозирования потребления электроэнергии в Бразилии, ЮАР, Японии, США, Тайване, Чехии [3, 12, 14-17] с достаточно высокой точностью, при этом, однако, необходимо отметить достаточно регулярный характер сигналов, описывающих энергопотребление в этих странах. Анализ рядов потребления электроэнергии в Украине показывает высокий уровень нестационарности сигналов, наличие резких выбросов и провалов, нерегулярных трендов и т.п., что не позволяет применить хорошо известные рекуррентные ИНС (Элмана, Джордана, Вильямса-Зипсера, Вольтерра и т.п.) в «чистом» виде и требует их модификации. В настоящей работе предпринимается попытка использования модифицированной рекуррентной сети Элмана [18] для решения задачи прогнозирования потребления электроэнергии в Украине на уровне региональных энергосистем.

Архитектура прогнозирующей рекуррентной нейронной сети. Архитектура модифицированной рекуррентной нейронной сети Элмана приведена на рис. 1. Ее использование предполагает, что процесс энергопотребления имитируется выходным сигналом некоторой нелинейной динамической системы, возмущаемой множеством факторов (погодных, временных и т.п.), включая и прошлые состояния системы. Элман предложил в дополнение к традиционным скрытым и выходному слоям ИНС ввести в сеть дополнительный слой обратной связи, именуемый контекстным, или слоем состояний. Этот слой получает сигналы с выхода скрытого слоя и через элементы задержки z^{-1} подает их на предыдущий (у Элмана входной) слой, сохраняя таким образом обрабатываемую информацию с предыдущих тактов внутри сети. Предлагаемая нами модификация касается первого и выходного слоев рекуррентной нейронной сети.

«Строительными блоками» рассматриваемой здесь архитектуры являются стандартные нейроны (элементарные персептроны Розенблатта) с активационными функциями типа сигмоиды, элементы задержки z^{-1} и блоки фаззификации, предназначенные для преобразования входных порядковых и номинальных переменных, характеризующих влияние окружающей среды, в количественную форму. Таким образом, в модифицированную ИНС вводится дополнительный первый скрытый слой, полностью совпадающий с первым слоем прогнозирующей NARX-нейро-фаззи-системы, предложенной в [19]. Кроме того, вместо адаптивного линейного ассоциатора в выходном слое опять-таки используется нелинейный нейрон с сигмоидой в качестве активационной функции, что улучшает экстраполирующие свойства сети.

Сигналы первого скрытого слоя задержек и фаззификации в виде $(n \times 1)$ -вектора $o^{[1]}(k) = (o_1^{[1]}(k), \dots, o_n^{[1]}(k))^T$ с компонентами, описывающими текущее потребление электроэнергии $y(k)$, его прошлые значения $y(k-1), \dots, y(k-d)$, временные и погодные характеристики, преобразованные в числовую форму блоками фаззификации, поступают на второй скрытый слой, образованный n_2 идентичными нейронами с сигмоидальными активационными функциями $\psi_j^{[2]}$, $j = 1, 2, \dots, n_2$ и $n_2(n_2 + 1)$ настраиваемыми синаптическими весами $w_{ji}^{[2]}$.

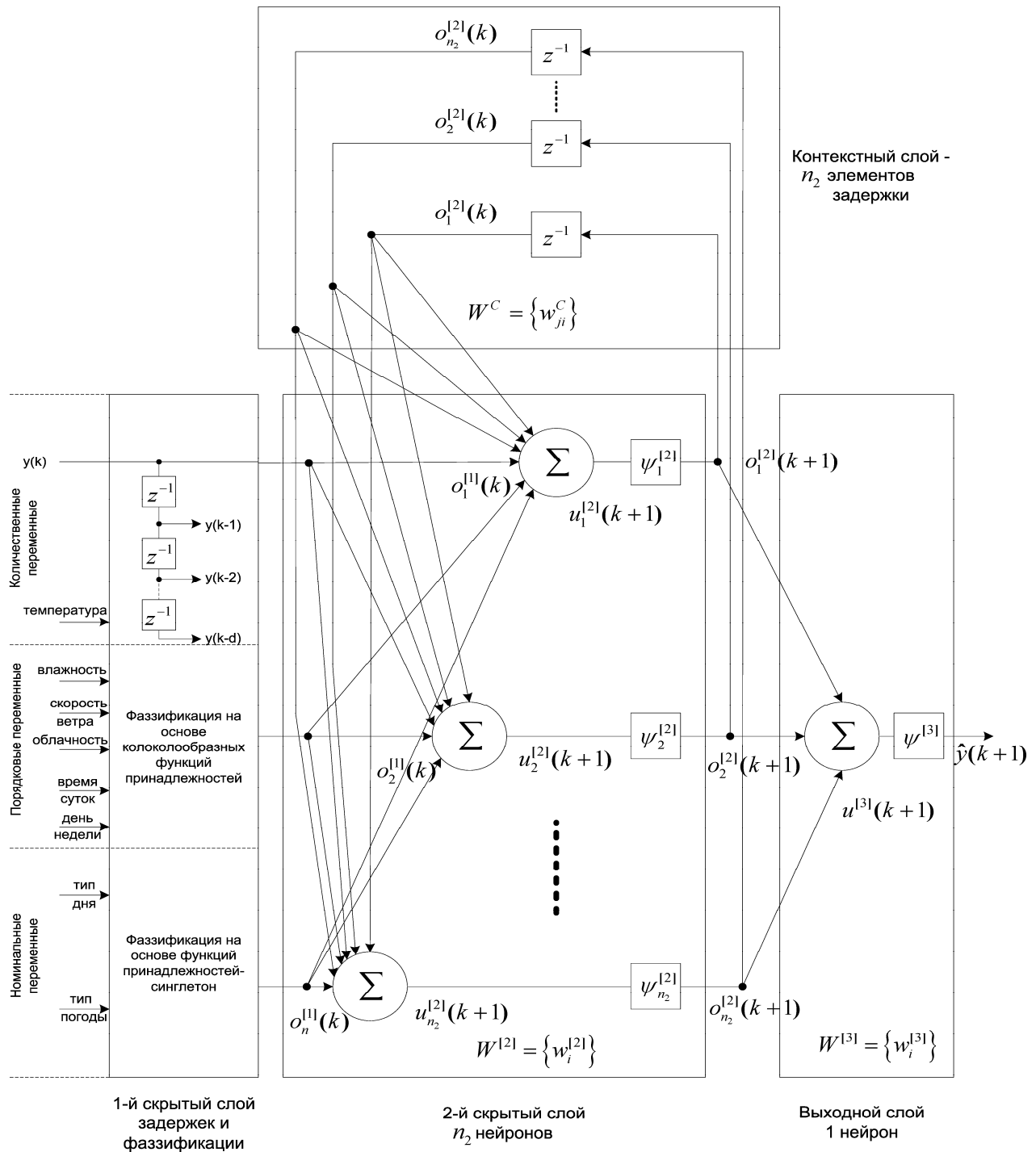


Рис. 1 – Модифицированная рекуррентная нейронная сеть Элмана

Как можно видеть из приведенной архитектуры, выходной сигнал j -го нейрона второго скрытого слоя может быть представлен в виде

$$\begin{cases} u_j^{[2]}(k+1) = \sum_{i=1}^n w_{ji}^{[2]} o_i^{[1]}(k) + \sum_{i=1}^{n_2} w_{ji}^C o_i^{[2]}(k) + \theta_j^{[2]}, \\ o_j^{[2]}(k+1) = \psi_j^{[2]}(u_j^{[2]}(k+1)), j = 1, 2, \dots, n_2 \end{cases} \quad (1)$$

(здесь $k = 0, 1, 2, \dots$ – текущее дискретное время, $\theta_j^{[2]}$ – сигнал смещения j -го нейрона второго скрытого слоя), а выход слоя в целом – в векторно-матричной форме

$$o^{[2]}(k+1) = \Psi^{[2]}(W^{[2]}x^{[1]}(k) + W^C o^{[2]}(k)), \quad (2)$$

где $o^{[2]}(k+1) - (n_2 \times 1)$ – векторный сигнал, который далее передается на выходной слой в форме $x^{[3]}(k+1) = (1, o^{[2]T}(k+1))^T$, $\Psi^{[2]} = \text{diag}\{\psi_j^{[2]}\} - (n_2 \times n_2)$ – диагональная матрица активационных функций, $W^{[2]} - (n_2 \times (n_2 + 1))$ – матрица настраиваемых синаптических весов второго скрытого слоя, $x^{[2]}(k) = (1, o^{[1]T}(k))^T - (n+1) \times 1$ – вектор входных сигналов второго скрытого слоя, поступающих из первого скрытого слоя, $W^C - (n_2 \times n_2)$ – матрица настраиваемых синаптических весов контекстного слоя.

Контекстный слой образован n_2 элементами задержки z^{-1} , которые обеспечивают чистое запаздывание выходных сигналов второго слоя, после чего задержанные значения $o_j^{[2]}(k)$ вновь подаются через синаптические веса w_{ji}^C на сумматоры нейронов второго скрытого слоя. Объединяя все входные сигналы второго скрытого слоя в составной вектор $x^{[2]}(k) = (1, o^{[1]T}(k), o^{[2]T}(k))^T = (x^{[1]T}(k), o^{[2]T}(k))^T$ размерности $(1+n+n_2) \times 1$, можно записать преобразование, реализуемое совместно вторым скрытым и контекстным слоями, в виде

$$o^{[2]}(k+1) = \Psi^{[2]}(W^{[2]C} x^{[2]}(k)), \quad (3)$$

где матрица настраиваемых синаптических весов имеет размерность $n_2 \times (1+n+n_2)$.

Выходной слой модифицированной рекуррентной нейронной сети образован единственным нейроном, формирующим скалярный сигнал прогноза

$$\hat{y}(k+1) = \psi^{[3]}(u^{[3]}(k+1)) = \psi^{[3]}(\sum_{i=0}^{n_2} w_i^{[3]} x_i^{[3]}(k+1)) = \psi^{[3]}(w^{[3]T} x^{[3]}(k+1)), \quad (4)$$

где $w^{[3]} - (n_2 + 1) \times 1$ – вектор настраиваемых синаптических весов выходного слоя.

В общем же виде преобразование, реализуемое предлагаемой модифицированной сетью Элмана, может быть записано в виде

$$\begin{cases} o^{[2]}(k+1) = \Psi^{[2]}(W^{[2]}x^{[1]}(k) + W^C o^{[2]}(k)) = \Psi^{[2]}(W^{[2]C} x^{[2]}(k)), \\ \hat{y}(k+1) = \psi^{[3]}(w^{[3]T} x^{[3]}(k+1)). \end{cases} \quad (5)$$

Обучение прогнозирующей рекуррентной нейронной сети. Обучение рассматриваемой нейронной сети будем проводить путем пошаговой минимизации стандартного локального квадратичного критерия

$$\begin{aligned} E(k+1) &= \frac{1}{2} e^2(k+1) = \frac{1}{2} (y(k+1) - \hat{y}(k+1))^2 = \\ &= \frac{1}{2} (y(k+1) - \psi^{[3]}(u^{[3]}(k+1)))^2 = \frac{1}{2} (y(k+1) - \psi^{[3]}(\sum_{i=0}^{n_2} w_i^{[3]} x_i^{[3]}(k+1)))^2, \end{aligned} \quad (6)$$

где

$$\psi^{[s]}(u^{[s]}(k)) = \frac{1}{1 + e^{-\gamma u^{[s]}(k)}}, \quad (7)$$

$\gamma > 0$ – параметр, определяющий форму активационной функции и также, в принципе, настраиваемый, $s = 2, 3$ – номер слоя.

Процесс обучения основывается на концепции обратного распространения ошибок и начинается с настройки синаптических весов единственного нейрона выходного слоя.

Для текущего момента времени k , когда доступна информация о значениях $y(k), e(k)$, можно записать градиентную процедуру минимизации (6) в виде

$$\begin{aligned} w_i^{[3]}(k+1) &= w_i^{[3]}(k) - \eta^{[3]}(k) \frac{\partial E(k)}{\partial e(k)} \square \frac{\partial e(k)}{\partial u^{[3]}(k)} \square \frac{\partial u^{[3]}(k)}{\partial w_i^{[3]}(k)} = \\ &= w_i^{[3]}(k) + \eta^{[3]}(k) e(k) \frac{\partial \psi^{[3]}(u^{[3]}(k))}{\partial u^{[3]}(k)} x_i^{[3]}(k) = \\ &= w_i^{[3]}(k) + \eta^{[3]}(k) \delta^{[3]}(k) x_i^{[3]}(k), i = 0, 1, \dots, n_2, \end{aligned} \quad (8)$$

где $\eta^{[3]}(k)$ – параметр шага поиска, выбираемый обычно из эмпирических соображений,

$\delta^{[3]}(k) = e(k) \frac{\partial \psi^{[3]}(u^{[3]}(k))}{\partial u^{[3]}(k)} = \frac{\partial E(k)}{\partial u^{[3]}(k)}$ – локальная ошибка (δ -ошибка) выходного слоя.

В компактной векторной форме алгоритм (8) можно переписать

$$w_i^{[3]}(k+1) = w_i^{[3]}(k) + \eta^{[3]}(k) \delta^{[3]}(k) x_i^{[3]}(k) = w^{[3]}(k) + \eta^{[3]}(k) \delta^{[3]}(k) \nabla_{w^{[3]}} u^{[3]}(k), \quad (9)$$

а с учетом (7) –

$$w^{[3]}(k+1) = w^{[3]}(k) + \eta^{[3]}(k) \gamma e(k) \hat{y}(k) (1 - \hat{y}(k)) x^{[3]}(k). \quad (10)$$

Соответствующим выбором параметра шага $\eta^{[3]}(k)$ можно попытаться увеличить скорость процесса обучения, переходя тем самым от градиентных к более изощренным процедурам минимизации.

Рассматривая одношаговую модификацию алгоритма обучения Левенберга-Марквардта [20]

$$\begin{aligned} w^{[3]}(k+1) &= w^{[3]}(k) + (\nabla_{w^{[3]}} u^{[3]}(k) \nabla_{w^{[3]}}^T u^{[3]}(k) + \rho I_{n_2+1})^{-1} \delta^{[3]}(k) \nabla_{w^{[3]}} u^{[3]}(k) = \\ &= w^{[3]}(k) + (x^{[3]}(k) x^{[3]T}(k) + \rho I_{n_2+1})^{-1} \delta^{[3]}(k) x^{[3]}(k) \end{aligned} \quad (11)$$

(здесь $\rho > 0$ – параметр регуляризации, $I_{n_2+1} - ((n_2 + 1) \times (n_2 + 1))$ – единичная матрица), после несложных преобразований, основанных либо на формуле Шермана-Моррисона, либо на псевдообращении по Муру-Пенроузу, можно переписать (11) в простом виде

$$w^{[3]}(k+1) = w^{[3]}(k) + \frac{\delta^{[3]}(k) x^{[3]}(k)}{\rho + \|x^{[3]}(k)\|^2}, \quad (12)$$

который структурно совпадает с аддитивной модификацией адаптивного алгоритма идентификации Качмажа [21], а при $\rho = 0$ – с алгоритмом обучения нейронных сетей Уидроу-Хоффа.

Элементарный анализ (10), (12) показывает, что процесс обучения резко замедляется в случае попадания на «хвосты» сигмолды, где значения $\hat{y}(k)$ близки к единице. В этом случае представляется целесообразным использование регуляризованных процедур обучения, к которым относится алгоритм Чана-Фоллсайда [22]

$$w^{[3]}(k+1) = w^{[3]}(k) + \eta^{[3]} \delta^{[3]}(k) x^{[3]}(k) + \rho \Delta w^{[3]}(k-1), \quad (13)$$

(здесь $\eta^{[3]} - \text{const} > 0$; $1 > \rho > 0$; $\Delta w^{[3]}(k-1) = w^{[3]}(k) - w^{[3]}(k-1)$), успешно проходящий плато целевой функции, порождаемые «хвостами» сигмолды.

Повысить скорость сходимости процедуры (13) можно, используя ее «гибрид» с (12) в виде [23]

$$w^{[3]}(k+1) = w^{[3]}(k) + \frac{\eta^{[3]} \delta^{[3]}(k) x^{[3]}(k) + \rho \Delta w^{[3]}(k-1)}{\|x^{[3]}(k)\|^2}, \quad (14)$$

обладающий более высоким быстродействием.

Алгоритм одновременной настройки синаптических весов второго скрытого и контекстного слоев может быть записан в виде

$$\begin{aligned} w_{ji}^{[2]C}(k+1) &= w_{ji}^{[2]C}(k) - \eta^{[2]C}(k) \frac{\partial E(k)}{\partial w_{ji}^{[2]C}(k)} = \\ &= w_{ji}^{[2]C}(k) - \eta^{[2]C}(k) \frac{\partial E(k)}{\partial u_j^{[2]}(k)} \square \frac{\partial u_j^{[2]}(k)}{\partial w_{ji}^{[2]C}(k)} = w_{ji}^{[2]C}(k) + \eta^{[2]C}(k) \delta_j^{[2]C}(k) x_i^{[2]}(k), \end{aligned} \quad (15)$$

где $\delta_j^{[2]C}(k) = \frac{\partial E(k)}{\partial u_j^{[2]}(k)}$, $j = 1, 2, \dots, n_2$; $i = 0, 1, \dots, n + n_2$.

Записывая локальную ошибку этих слоев в виде

$$\begin{aligned} \delta_j^{[2]C}(k) &= \frac{\partial E(k)}{\partial u_j^{[2]}(k)} = \frac{\partial E(k)}{\partial o_j^{[2]}(k)} \square \frac{\partial E(k)}{\partial o_j^{[2]}(k)} = \frac{\partial E(k)}{\partial u^{[3]}(k)} \square \frac{\partial u^{[3]}(k)}{\partial o_j^{[2]}(k)} = \\ &= \frac{\partial E(k)}{\partial u^{[3]}(k)} \square \frac{\partial}{\partial o_j^{[2]}(k)} \sum_{p=0}^{n_2} w_p^{[3]}(k) x_p^{[3]}(k) = \delta^{[3]}(k) w_j^{[3]}(k), \end{aligned} \quad (16)$$

с учетом того, что

$$o_j^{[2]} = \psi_j^{[2]}(u_j^{[2]}(k)), \quad (17)$$

несложно получить

$$\delta_j^{[2]C}(k) = \frac{\partial E(k)}{\partial o_j^{[2]}(k)} \square \frac{\partial \psi_j^{[2]}(u_j^{[2]}(k))}{\partial u_j^{[2]}(k)}. \quad (18)$$

Представив далее $\frac{\partial E(k)}{\partial o_j^{[2]}(k)}$ в форме

$$\begin{aligned} \frac{\partial E(k)}{\partial o_j^{[2]}(k)} &= \frac{\partial E(k)}{\partial u^{[3]}(k)} \square \frac{\partial u^{[3]}(k)}{\partial o_j^{[2]}(k)} = \\ &= \frac{\partial E(k)}{\partial u^{[3]}(k)} \square \frac{\partial}{\partial o_j^{[2]}(k)} \sum_{p=0}^{n_2} w_p^{[3]}(k) x_p^{[3]}(k) = \delta^{[3]}(k) w_j^{[3]}(k), \end{aligned} \quad (19)$$

можно переписать (18)

$$\delta_j^{[2]C}(k) = \frac{\partial \psi_j^{[2]}(u_j^{[2]}(k))}{\partial u_j^{[2]}(k)} \delta^{[3]}(k) w_j^{[3]}(k), \quad (20)$$

откуда следует

$$\begin{aligned} w_{ji}^{[2]C}(k+1) &= w_{ji}^{[2]C}(k) + \eta^{[2]C}(k) x_i^{[2]}(k) \frac{\partial \psi_j^{[2]}(u_j^{[2]}(k))}{\partial u_j^{[2]}(k)} \delta^{[3]}(k) w_j^{[3]}(k) = \\ &= w_{ji}^{[2]C}(k) + \eta^{[2]C}(k) x_i^{[2]}(k) \delta_j^{[2]}(k), \end{aligned} \quad (21)$$

где $\delta_j^{[2]C}(k) = \frac{\partial \psi_j^{[2]}(u_j^{[2]}(k))}{\partial u_j^{[2]}(k)} \delta^{[3]}(k) w_j^{[3]}(k)$.

Таким образом, процесс обучения нейронов второго скрытого слоя с дополнительными входами из контекстного слоя в общем случае может быть записан в виде

$$w_j^{[2]C}(k+1) = w_j^{[2]C}(k) + \eta^{[2]C}(k) \delta_j^{[2]C}(k) x^{[2]}(k), \quad (22)$$

а с учетом (14) –

$$w_j^{[2]C}(k+1) = w_j^{[2]C}(k) + \frac{\eta^{[2]C} \delta_j^{[2]C}(k) x^{[2]}(k) + \rho \Delta w_j^{[2]C}(k-1)}{\|x^{[2]}(k)\|^2}. \quad (23)$$

В процессе обучения модифицированной рекуррентной нейронной сети сначала необходимо последовательно вычислить локальные ошибки $\delta^{[3]}(k), \delta_j^{[2]C}(k)$, параметры шага $\eta^{[3]}(k), \eta^{[2]C}(k)$, а затем уже переходить собственно к настройке синаптических весов.

Выводы

В работе предложена модифицированная архитектура рекуррентной прогнозирующих нейронных сетей и алгоритмы ее обучения для оперативного прогнозирования потребления электрической энергии. Простота и быстрота предложенной нейросети обеспечивает ей преимущество по сравнению с традиционными подходами, используемыми ныне для решения этой задачи.

Литература

1. Comparative Models for Electrical Load Forecasting / Ed. by Bunn D.W., Farmer E.D. – New York: John Wiley&Sons, 1985.
2. Bunn D.W. Forecasting loads and prices in competitive power markets // Proc. IEEE. – 2000. – 88. – P. 163-169.
3. Hippert H.S., Pedreira C.E., Souza R.C. Neural networks for short-term load forecasting: a review and evaluation // IEEE Trans. Power Systems. – 2001. – 16. – N. 1. – P. 44-55.
4. Ткаченко Р.О., Павлюк О.М. Прогнозування споживання електричної енергії у Львівській області за допомогою штучних нейронних мереж // Вісник НУ «Львівська політехніка». Комп'ютерна інженерія та інформаційні технології. – 2002. – 450. – С. 76-80.
5. Бардачев Ю.Н., Гринавцев О.В., Литвиненко В.И., Фелелов А.А. Синтез и анализ работы нечетких нейронных сетей с помощью иммунных алгоритмов при прогнозировании электрической нагрузки // Моделирование та керування станом еколого-економічних систем регіону. – 2006. – 3. – С. 47-68.
6. Прихно В.Л., Черненко П.А. Модели, методы и программные средства для анализа и прогнозирования электрических нагрузок и для решения задач оперативного управления ЭЭС // Сб. науч. тр. ИЭД НАНУ. – К.: ИЭД НАН Украины, 2007. – N. 3(18). – С. 26-33.
7. Yuan J.-L., Fine T.L. Forecasting demand for electric power // Advances in Neural Information Processing Systems / Ed. by S.J. Hanson, J.D. Cowan, C.L. Giles – San Mateo, CA: Morgan Kauffman, 1993. – 5. – P. 739-746.
8. Connor J.T. A Robust Neural Network Filter for Electricity Demand Prediction // Journal of Forecasting. – 1996. – 15. – N. 6. – P. 437-458.
9. Piras A., Germond A., Buchenel B., Imhof K., Jaccard Y. Heterogeneous artificial neural network for short term electrical load forecasting // IEEE Trans. Power Systems. – 1996. – 11. – N. 2. – P. 397-402.
10. Tzafestas S., Tzafestas E. Computational intelligence techniques for short-term electric load forecasting // Journal of Intelligent and Robotic Systems. – 2001. – 31. – P. 7-68.
11. Ling S.H., Leung F.H.F., Lam H.K., Tam P.K.S. Short-term electric load forecasting based on a neural fuzzy network // IEEE Trans. Industrial Electronics. – 2003. – 50. – N. 6. – P. 1305-1316.
12. Vermaak J., Both E.C. Recurrent Neural Networks for Short-Term Load Forecasting // IEEE Trans. Power Systems. – 1998. – 13. – N. 1. – P. 126-132.
13. Mandic D.P., Chambers J.A. Recurrent Neural Networks for Prediction. – Chichester: John Wiley&Sons, 2001. – 285 p.
14. Mori H., Hidenori K. Optimal fuzzy inference for short-term load forecasting // IEEE Trans. Power Systems. – 1996. – 11. – N. 1. – P. 390-396.
15. Khotanzad A., Afkhami-Rohani R., Lu T.L., Abaye A., Davis M., Maratukulam D. ANNSTLF – A Neural-Network-Based Electric Load Forecasting System // IEEE Trans. Neural Networks. – 1997. – 8. – N. 4. – P. 835-846.

16. Khan M.R., Abraham A. A Hybrid Fuzzy-Neural Network for Modelling Short-Term Demand Forecasting in Czech Republic // Second International Workshop on Intelligent Systems Design and Applications. – Atlanta, USA, 2002. – 187-194.
17. Pai P.-F., Hong W.C. Forecasting regional electricity load based on recurrent support vector machines with genetic algorithms // Electric Power Systems Research. – 2005. – 74. – P. 417-425.
18. Elman J.L. Finding structure in time // Cognitive Science. – 1990. – 14. – P. 179-211.
19. Bodyanskiy Y., Popov S., Rybalchenko T. Multilayer neuro-fuzzy network for short term electric load forecasting // Lecture Notes in Computer Science – Berlin, Heidelberg: Springer-Verlag, 2008. – 5010. – P. 339-348.
20. Shepherd A.J. Second-Order Methods for Neural Networks (Fast and Reliable Training Methods for Multi-Layer Perceptrons). – London: Springer-Verlag, 1997. – 145 p.
21. Райбман Н.С., Чадеев В.М. Построение моделей процессов производства. – М.: Энергия, 1975. – 376 с.
22. Cichocki A., Unbehauen R. Neural Networks for Optimization and Signal Processing. – Stuttgart: Teubner, 1993. – 526 p.
23. Бодянский Е.В., Викторов Е.А. Каскадная ортогональная нейронная сеть на двойных ортонах и алгоритм ее обучения в задачах обработки информации // Доклады XV международной конференции по автоматическому управлению "Автоматика – 2008". – Одесса, 2008. – С. 70-73.

МОДИФІКОВАНА РЕКУРЕНТНА НЕЙРОННА МЕРЕЖА ЕЛМАНА В ЗАДАЧІ КОРОТКОТЕРМІНОВОГО СПОЖИВАННЯ ЕЛЕКТРОЕНЕРГІЇ

І.П. Плісс, С.В. Попов, Т.В. Рибальченко, О.В. Сліпченко

Запропоновано модифіковану архітектуру рекурентної нейронної мережі Елмана, яка спроможна обробляти дані, задані в різних шкалах вимірювання. Розроблено відповідні алгоритми навчання, засновані на процедурах 1-го та 2-го порядків, а також захищені від «паралічу» мережі.

MODIFIED ELMAN RECURRENT NEURAL NETWORK FOR SHORT-TERM ELECTRIC LOAD FORECASTING

I.P. Pliss, S.V. Popov, T.V. Rybalchenko, E.V. Slipchenko

Modified architecture of Elman recurrent neural network is proposed that is capable to process data presented in different measurement scaled. Learning algorithms are developed, based on first and second order procedures with network paralysis-proof modifications.